A Study on Deep Learning Based Packet Transmission Strategy for Intelligent Network Traffic Control

著者	毛 伯敏
number	63
学位授与機関	Tohoku University
学位授与番号	情博第689号
URL	http://hdl.handle.net/10097/00127262

氏名	毛 伯 敏 (MAO BO MIN)
学 位 の 種 類	博　士（情報科学）
学 位 記 番 号	情　博　第　689　号
学位授与年月日	平成３１年　３月２７日
学位授与の要件	学位規則第４条第１項該当
研 究 科 、 専 攻	東北大学大学院情報科学研究科（博士課程）応用情報科学専攻
学 位 論 文 題 目	A Study on Deep Learning Based Packet Transmission Strategy for Intelligent Network Traffic Control （深層学習に基づくパケット伝送ストラテジーを用いた知的ネットワーク制御に関する研究）
論 文 審 査 委 員	（主査）東北大学教　授　加藤 寧 東北大学教　授　篠原 歩　　東北大学教　授　張山　昌論 東北大学准教授　ファドゥルラ　ズバイル

論 文 内 容 の 要 旨

Chapter 1 Introduction

In recent years, the network traffic is increasing tremendously. The global Internet Protocol (IP) traffic per annum exceeded the ZettaByte (ZB) threshold at the end of 2016, and is expected to increase up to 2.3 ZB by 2020. At the same time, the number of devices connected to the IP networks will be three times as high as the global population in 2020. The packets sent by the devices have different requirements for the Quality of Service (QoS) since these devices belong to heterogeneous networks. Thus, it becomes increasingly crucial but complicated to control the network traffic when designing the routing strategy. Conventional rule-based routing strategies (e.g., the Shortest Path (SP) algorithm and so forth) which choose the paths according to the maximum or minimum metric values are usually attributed to the shortcomings of a significantly slow convergence. Furthermore, they may not be particularly suited to the multi-metric networks because the relationships among multiple metrics are difficult to be manually estimated. In order to exploit the complex relationships among the various metrics to decide the best path, machine learning-based intelligent network traffic control systems have drawn much attention in a wide spectrum of network environments. Supervised, unsupervised, and reinforcement learning techniques have been exploited to manage the packet routing in many different network scenarios, such as Wireless Mesh Networks (WMNs). However, these intelligent strategies are still based on the traditional rule-based routing due to the inefficiency of conventional machine learning techniques in dealing with multiple network parameters as well as the difficulties of characterizations of the inputs and outputs.

Nowadays, deep learning, a new breed of machine learning technique, has been widely applied in various fields, such as image classification, pattern recognition, and natural language processing. This technique can be utilized to effectively analyze the complex relationships among multiple inputs through training with example data. The trained deep learning architecture can predict the values of some parameters when we input the necessary information. Since the deep learning technique has exhibited superior performance in extremely difficult applications which have traditionally been dominated by humans, e.g. board games, it may also have interesting applications for network traffic control.

Inspired by the development of deep learning and the computation hardware, the Artificial Intelligence (AI) technique has been regarded as one of most important technologies to improve users' experience. Inspired by the flexibility and accuracy of deep learning, researchers have made many attempts to adopt this technology to optimize the network performance. To alleviate the increasing traffic overhead, this dissertation considers the deep learning technique to predict the routing paths. Since the deep learning technique consists of so many architectures and three training manners, this dissertation discusses the architecture design for different network scenarios, especially about the characterization of input and output. Moreover, we will also analyze the computation overhead of the deep learning based packet transmission strategies and propose novel computation platforms to conduct the algorithms. Even though the deep learning technique concerns more computation overhead compared with conventional routing algorithms, we will propose different platforms and deployment manners corresponding to the considered intelligent routing algorithms, which can significantly reduce the computation time. Furthermore, in this dissertation, we will focus on not only the static core networks, but also the dynamic networks considering link failures. The theoretical analysis and simulation will be conducted to evaluate the performance of the proposed deep learning architectures.

Chapter 2 Overview of Deep Learning and Traffic Control

As we mentioned in the Chapter 1, the global networks are confronted by increasing traffic overhead and growing complexity. Since it has been illustrated the superiority over human beings in complex activities, such as the games, image classification, and speech recognition, the technology of deep learning is promising to alleviate the traffic overhead. Before discussing the proposed deep learning based traffic control strategies, it is necessary to introduce some preliminaries of this new technique. We will study the mainly concerned calculations as well as the three training manners. After that, several commonly utilized deep learning architectures are discussed, of which the related applications in the networking field are explained to evaluate the advantages of deep learning in performance optimization. From the discussion, it can be clearly found that the deep learning technique is very flexible and efficient, which enables its wide perspectives in network performance optimization. Since our purpose is to address the traffic control challenge, we also survey the traditional strategies in different layers. We can conclude that the traffic control can be cooperatively conducted by different layers. However, to begin our research, this dissertation focuses on the routing design with deep learning to improve the traffic control performance. Then, we analyze the research contents of adopting the emerging AI technology to alleviate the traffic challenge in current networks.

Chapter 3 Deep Learning Based Routing Algorithm for Core Networks Running on GPU accelerated SDRs

To adopt the deep learning technique for improving the traffic control, we first need to choose a network scenario in this chapter. Since the routing is concerned with packet forwarding in the Internet, we choose the static backbone network as our considered scenario in this chapter. As we mentioned in Chapter 1, the proprietary hardware architectures lack the flexibility for the potential update of management strategies. To apply the state-of-the-art software driven routing algorithms developed for different network services, it is necessary to improve the programmability of the core routers. Moreover, since the deep learning technique is concerned with massive matrix computations, we consider the GPU-accelerated Software Defined Routers (SDRs) as the routing architecture. Specifically, we adopt the collected data comprising inbound traffic patterns and corresponding subsequent nodes (i.e., routers) to train the Deep Belief Architectures (DBAs) in a supervised manner. Then, the trained DBAs can compute the subsequent nodes (i.e., routers) with the traffic patterns of the edge routers as the input. And the deep learning related training and running operations are executed by the GPU-accelerated SDRs to expedite the computation of the intelligent protocol. Furthermore, the packet forwarding operations as well as the processing work are cooperatively conducted by the GPUs and CPUs. The numerical analysis illustrates that the GPU resource can significantly reduce the time consumption. Moreover, the simulation results demonstrate that the proposed deep learning based routing can achieve much better network performance compared with conventional routing protocols.

Chapter 4 Online Learning Based Routing Strategy for Software Defined Communication Systems

In last chapter, the supervised learning based routing strategy is proposed to tackle the increasing traffic overhead in the backbone networks. Since the labeled data impact on the performance of the considered deep learning architectures, it is a critical step to collect the training data. However, in many scenarios, it is very difficult to collect enough satisfying labeled data, for which the heterogeneous network is a good example. In this chapter, we consider the Software Defined Communication Systems (SDCSs), of which the data planes consist of various communication technologies, leading to fast changing traffic patterns. To address the limitation in the collection of enough training data, in this chapter, we propose an online learning based routing strategy which periodically trains the considered architecture with real-time traffic patterns. The proposal consists of two steps: the initial phase and running phase. In the initial phase, the controller runs the conventional routing protocol while the switches record the traffic trace, which is utilized by the controller to initialize the utilized CNNs. Then, in the running phase, the CNNs are adopted in the controller to choose paths. Furthermore, to adapt the trained CNNs to the changing traffic patterns as well as reduce the training computation overhead, in the running phase, the switches keep recording the traffic trace for periodically retraining the CNNs in the controller. In this way, the CNNs can get adaptive to the changing traffic patterns due to the periodical training with newly collected traffic trace. Additionally, the simulation analysis demonstrates the advantages of the online learning method and the intelligent routing method is illustrated to outperform the conventional strategy in terms of network performance.

Chapter 5 Value Iteration Based Deep Learning Architecture for Routing in Dynamic Networks

In previous two chapters, we focus on the static networks. However, current global networks consist of so many dynamic parts. In this chapter, we discuss the deep learning based packet forwarding for dynamic networks. To predict the routing paths for the dynamic networks, the proposed intelligent strategies are not available since the considered input just considers the node information. And according to our descriptions in above chapters, it can be easily found that the deep learning architecture design is related to the network topology. Therefore, to fit the topology changes, we propose a Value Iteration Architecture (VIA) which takes the node information as well as the whole network topology as the input. The considered reinforcement training manner enables the VIA to learn the routing policy independent on the network topology. Therefore, once given the network scenario, the trained VIA can be utilized to predict the paths directly. Moreover, this chapter discusses the time complexity of the proposal. Besides the network performance improvement, the proposed Heterogeneous Computation Platform (HCP) can accelerate the execution of the proposal and the considered deployment manner can further reduce the computation overhead.

Chapter 6 Conclusion

In this dissertation, we study the deep learning technique to tackle the increasing traffic overhead in global networks. We consider three different network scenarios and discuss their traffic patterns. To predict the routing paths with the deep learning technology, we propose three architectures, which can efficiently alleviate the traffic overhead. It can be concluded that the deep learning technique can be utilized to efficiently tackle the challenges of globally increasing traffic. The various deep learning architectures and different training manners significantly increase the flexibility, leading to great potential to be applied in practical network deployment. To further improve the network performance, more meaningful research can be conducted in the future.

論文審査結果の要旨

既存のグローバルネットワークは、近年トラヒックのオーバーヘッドの増大とネットワークの複雑さという問題に直面している。トラヒックのオーバーヘッドを軽減するために、本論文では新たなパケット伝送方式を提案している。本研究では、ネットワークトラヒック制御の主な課題を分析し、深層学習の運用方法について議論し、様々なネットワークシナリオについて具体的な提案を行っている。本論文は6つの章により構成されている。

第1章では、本研究の背景を説明している。深層学習とハードウェア開発のブレークスルーについて議論した後、本研究の目的と内容について述べている。

第2章では、主に深層学習技術と従来のネットワーク制御ソリューションについて説明している。まず、深層学習技術の基礎とネットワーキングにおける応用について述べ、次に従来のトラヒック制御手法を分析し、それぞれの特徴により分類している。

第3章では、静的なバックボーンネットワークのための深層学習に基づくルーティング戦略を提案している。 Deep Belief Architectures（DBA）は、次ノードを予測するために導入されている。インテリジェントルーティング戦略を効率的に実行するために、GPU アクセラレーションルータアーキテクチャを検討し、パケット処理ステップの詳細な分析を行っている。その結果、GPU を利用した場合、深層学習がトラヒック制御性能を改善し、計算速度が向上できることを確認している。この点は、実用上の観点から高く評価できる。

第4章では、蓄積が難しい訓練データをより効率的に蓄積するため、深層学習アーキテクチャを用いたオンライン学習方法について提案している。シミュレーションを用いて、連続的なパフォーマンスの向上を評価し、提案アルゴリズムはトラヒックの変化に応じて最適化を行うことが可能であることを確認した。これは実時間利用の可能性を示すものであり、高く評価できる。

第5章では、動的ネットワークの経路を予測するための強化学習を用いた提案手法を示している。ヘテロジニアスコンピューティングプラットフォームは、より効果的なルーティングを実行するために利用されている。提案手法は動的ネットワークにおけるトラヒック制御に利用可能であり、トポロジが変化したときのパスを正確に予測できる点は十分に評価できる。

第6章では、本論文のまとめを行っている。

以上要するに本論文は、ネットワークトラヒック制御を改善するための深層学習を用いたパケット伝達方式を提案しているものであり、応用情報科学並びに情報通信技術の発展に寄与するところが少なくない。

よって、本論文は博士（情報科学）の学位論文として合格と認める。